

## Utilización de software libre para el análisis de la evolución de usos del suelo en la Cuenca del Segura mediante teledetección.

<sup>(1)</sup>F. Cánovas García, <sup>(2)</sup>J. Moreno Brotóns, <sup>(3)</sup>F.J. Gomariz Castillo y <sup>(4)</sup>F. Alonso Sarriá

Instituto Universitario del Agua y del Medio Ambiente (INUAMA), Universidad de Murcia, Campus de Espinardo, Edificio B. 30100 Murcia. <sup>(1)</sup>fulgencio.canovas@um.es, <sup>(2)</sup>jesusmb@um.es, <sup>(3)</sup>fjgomariz@um.es, <sup>(4)</sup>alonsarp@um.es

### RESUMEN

*En este trabajo se describe la utilización de herramientas de software libre, básicamente GRASS y R, para obtener una serie de mapas de coberturas del suelo (1976-2006) a partir de imágenes de satélite Landsat MSS y Landsat TM. Se trata de un proyecto concedido a un año, por lo que se requería una metodología que permitiera realizar el análisis de forma rápida y sencilla, aún tratando de aplicar técnicas de clasificación avanzadas. Dada la complejidad del trabajo y la premura de tiempo, se ha tratado de automatizar gran parte del trabajo mediante diversos scripts con BASH y R.*

*El procedimiento de clasificación utilizado ha sido la clasificación supervisada de máxima verosimilitud, utilizando como áreas de entrenamiento polígonos en los que la cartografía de usos y fotografía aérea mostraba un uso constante desde 1977 hasta la actualidad. Se ha tratado de mejorar el procedimiento mediante la utilización de dos imágenes por año, información textural y combinación de clasificación supervisada y no supervisada.*

**Palabras clave:** Teledetección, Usos del Suelo, Landsat

### ABSTRACT

*This work describes the utilization of free software tools, basically GRASS and R, to obtain a series of land cover maps (1976-2006) from satellite images of Landsat MSS and Landsat TM. It's a one year project, so a quick and simple methodology was needed, still trying to apply advanced techniques of classification. Because of the complexity of the work and the time constraints, most of the workload had to be automatized with BASH and R scripts. The initial information is a set of images from the Instituto Universitario del Agua y del Medio Ambiente database at the University of Murcia, complemented with images acquired specifically for this project.*

*The classification method used was the supervised classification of maximum likelihood, using polygons as training areas where the land cover cartography and aerial photography shown a constant cover since 1977 up to the present. The procedure has been tried to improve by using two images per year, when it has been possible, textural information and the combination of supervised and unsupervised classification.*

**Keywords:** Remote sensing, land cover, Landsat

## INTRODUCCIÓN

El objetivo de este trabajo es presentar la metodología aplicada en un entorno GNU/Linux para la elaboración de una colección multitemporal de coberturas del suelo en el ámbito de la Demarcación Hidrográfica del Segura (DHS) desde 1976 hasta 2006.

Esta primera versión de la colección de mapas ha sido desarrollada según las siguientes fases:

- Creación de una base de datos de áreas de entrenamiento para la clasificación de las imágenes.
- Corrección geométrica de las imágenes de satélite y posterior verificación de los resultados.
- Corrección atmosférica y de iluminación de la serie de imágenes disponibles.
- Desarrollo y aplicación de una metodología basada en técnicas estándar para generar mapas de coberturas del suelo a partir de las imágenes de satélite pre-procesadas.
- Validación cualitativa y cuantitativa de los resultados.

El trabajo se enmarca en un proyecto más amplio y ambicioso que trata de analizar la evolución espacio-temporal de los usos del suelo en el ámbito de la Demarcación Hidrográfica del Segura desde 1976 hasta la actualidad con técnicas de teledetección. Forma parte de un convenio de colaboración entre el INUAMA (Universidad de Murcia) y la Fundación Instituto Euromediterráneo del Agua (IEA).

Los primeros trabajos realizados en el marco de este convenio han sido los que trataban de conseguir una metodología unificada en entorno GNU/Linux que permitiera el procesamiento de la ingente cantidad de información espacial disponible. Piénsese que para abarcar el ámbito de la Demarcación Hidrográfica del Segura (algo menos de 19.000 Km<sup>2</sup>) son necesarias más de dos escenas de Landsat-5 y 4 escenas en el caso de los antiguos Landsat MSS. Además, la mayoría de los mapas de coberturas se han calculado con dos fechas diferentes (normalmente primavera y verano) con el objeto de mejorar la discriminación de cultivos. En total se han procesado más de 100 escenas del satélite Landsat (tabla 1).

A partir del desarrollo de la metodología presentada en esta comunicación la intención del equipo de trabajo ha sido calcular una primera versión de estos mapas de cobertura de usos, 32 en total, que permitiese comprobar la adecuación de los *scripts* programados. Se atiende más en esta comunicación al software empleado y al planteamiento general de los *scripts* que a los resultados en sí de la clasificación obtenida.

Tabla 1: Escenas pre-procesadas y clasificadas(EF= escena flotante).

Año	Sensor	Escena (día/mes/track/frame)
1976	MSS	13/01/214/34; 13/01/214/33; 14/01/215/34; 14/01/215/33; 29/07/214/33; 29/07/214/34; 30/07/215/33; 30/07/215/34
1977	MSS	27/06/214/33; 27/06/214/34
1978	MSS	24/03/214/33; 24/03/214/34; 10/07/214/33; 10/07/214/34; 13/11/214/33; 13/11/214/33; 14/11/215/33; 14/11/215/33
1979	MSS	10/08/214/33; 10/08/214/34; 11/08/215/33; 11/08/215/34
1980	MSS	04/08/214/33; 04/08/214/34; 05/08/215/33; 05/08/215/34
1981	MSS	31/01/214/33; 31/01/214/34; 01/02/215/33; 01/02/215/34; 30/07/214/33; 30/07/214/34; 31/07/215/33; 31/07/215/34
1982	MSS	08/01/214/33; 08/01/214/34
1983	MSS	14/01/199/33; 14/01/199/34; 09/07/199/33; 09/07/199/33; 16/07/200/33; 16/07/200/34
1984	MSS	17/01/199/33; 17/01/199/34; 11/07/199/33; 11/07/199/34
1985	MSS	27/01/199/33; 27/01/199/34; 14/07/199/33; 14/07/199/34
1986	MSS	14/01/199/33; 14/01/199/34; 18/08/199/33; 18/08/199/34
1987	TM	13/08/199/33; 13/08/199/34
1988	TM	24/03/199/33; 24/03/199/34; 15/08/199/EF
1989	TM	17/07/199/EF; 21/10/199/EF
1990	TM	05/08/199/EF
1991	TM	24/08/199/EF
1992	TM	20/04/199/EF; 10/08/199/EF
1993	TM	28/07/199/EF
1994	TM	29/06/199/EF
1995	TM	13/04/199/EF
1996	TM	11/02/EF
1997	TM	17/03/199/EF; 02/04/199/EF; 07/07/199/EF
1998	TM	20/03/199/33; 20/03/199/34; 11/08/199/EF
1999	TM	08/04/199/33; 07/09/199/EF
2000	TM	01/03/199/EF; 08/08/199/EF
2001	TM	13/03/199/33; 3/03/199/34; 26/07/199/EF; 01/12/199/EF
2002	TM	23/03/199/EF; 26/05/199/EF
2003	TM	10/03/199/33; 10/03/199/34; 29/05/199/33; 29/05/199/34
2004	TM	04/03/199/33; 15/11/199/33; 15/11/199/34
2005	TM	30/03/200/33; 10/05/199/34; 18/06/199/33; 26/06/199/33; 24/10/199/33; 24/10/199/34
2006	TM	16/07/199/EF
2007	TM	13/03/199/EF

## METODOLOGÍA

La base metodológica del proyecto son un conjunto de técnicas de clasificación de imágenes de satélite:

- Clasificación supervisada.
- Clasificación no supervisada.
- Incorporación de información textural y estacional.
- Asignación de píxeles a clases por máxima verosimilitud.

El programa utilizado para el tratamiento digital de las imágenes ha sido GRASS debido a la experiencia del equipo de trabajo con este software y a la potencia que

tiene a la hora de automatizar tareas y aplicar procedimientos de trabajo en forma de *script*.

La alta integración con R permite utilizar este último programa para implementar métodos sofisticados de análisis de datos y gráficos avanzados a partir de los datos de GRASS. En concreto ha servido para aplicar métodos de análisis estadístico de la información extraída de las áreas de entrenamiento.

A continuación se expone paso por paso las diferentes metodologías empleadas junto a los *scripts* que se han programado.

### **Georreferenciación y corrección geométrica**

El primer paso en el tratamiento de las imágenes de satélite ha sido su georreferenciación, que se ha basado en la búsqueda de un conjunto de puntos de control con los que calibrar una ecuación de transformación de tipo polinómico.

Se han localizado puntos de control homogéneamente repartidos para evitar sobreabundancia en un área, ya que los errores se incrementarían en las áreas submuestreadas.

Finalmente se ha optado por contar con una muestra inicial de 35 puntos para escenas del sensor MSS, 50 para las del sensor TM y 120 cuando se trataba de escenas flotantes del sensor TM.

Se ha observado que el error medio cuadrático no disminuye de forma significativa al aumentar el grado de los polinomios de transformación, pero si se incrementa el riesgo de que aparezcan distorsiones locales importantes. Por ello finalmente se ha optado por emplear transformaciones lineales.

GRASS dispone de un módulo específico para la localización de puntos de control (*i.points*) que permite localizar puntos en la imagen y en una capa de información de referencia ya georreferenciada. Este módulo proporciona el error cuadrático medio y los errores por punto. Como criterio general, se ha tratado de obtener valores de error cuadrático entorno al tamaño del píxel. En todo caso hay que tener en cuenta la presencia de distorsiones localmente elevadas, aunque la media sea reducida. Este hecho se ha constatado especialmente en las áreas costeras.

### **Cálculo de reflectividad: corrección atmosférica y de iluminación**

En el caso de un proyecto multitemporal como el que aquí se presenta, la corrección atmosférica resulta obligada. Se ha utilizado un método sencillo de corrección atmosférica y de iluminación que se describe en [1]. Para realizar esta tarea se ha programado un script completo que implementa la metodología desarrollada y que se resume a continuación.

Se ha realizado la corrección atmosférica según el método de *Chávez* [2] para todas las imágenes. De este modo la distorsión atmosférica residual que quede tras la corrección afectará, en cada imagen, del mismo modo a los píxeles de las áreas de entrenamiento utilizadas para caracterizar las clases y a los posteriormente clasificados minimizando su impacto en el error de clasificación.

Este enfoque obliga a disponer de áreas de entrenamiento válidas para todas las fechas, es decir que no hayan variado su uso a lo largo del tiempo. Para ello se ha optado por un método de búsqueda de estas áreas que se expondrá en el apartado correspondiente.

Con respecto a la corrección de iluminación se ha optado por un ajuste semiempírico como el realizado en el método lambertiano de corrección C [3] que aparece en la ecuación 1.

$$\rho_{k,i} = \rho_k \left( \frac{\cos\theta_i + c_k}{\cos\beta_i + c_k} \right) \quad (1)$$

donde  $\rho_k$  es la reflectividad de la banda  $k$ ,  $\theta_i$  es el ángulo cenital solar,  $\beta_i$  es el ángulo formado por la radiación incidente con el plano del terreno y  $c_k$  es una constante empírica para cada banda que se obtiene de:

$$c_k = \frac{b_k}{a_k} \quad (2)$$

donde  $b_k$  y  $a_k$  son la constante y la pendiente de la recta de regresión  $p_k = b_k + a_k \beta$ . Este método tiene la ventaja de que obtiene correcciones de iluminación diferentes para cada una de las bandas.

En la figura 1 se observa el descenso en los valores de reflectividad debido a la eliminación de la radiación atmosférica difusa en el proceso de corrección atmosférica. También muestra como la corrección de iluminación ha modificado los valores de reflectividad a ambos lados de la Sierra de Carrascoy. Se observa como en la vertiente de solana estos han disminuido mientras que en la de umbría han aumentado.

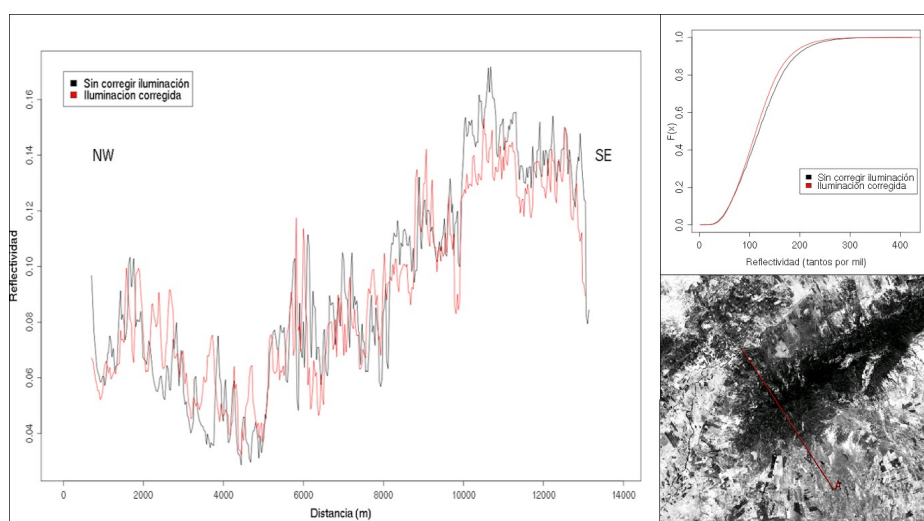


Figura 1: Resultado de la corrección atmosférica (superior derecha) en una escena y de iluminación en un transecto de la Sierra de Carrascoy (izquierda e inferior derecha).

### Creación de capas de información textural

Una de las tendencias más fructíferas en teledetección en los últimos años ha sido la incorporación de información adicional a las capas de reflectividad en los procedimientos de clasificación. Principalmente se trata de dos tipos: información contextual e información textural.

En este trabajo se ha incorporado la segunda, que representa la variación espacial del brillo de una imagen [4]. Según estos mismos autores, la utilización de este tipo de información resulta especialmente interesante en paisajes mediterráneos debido a la gran variedad y fragmentación de patrones espaciales.

Si las características espectrales del conjunto de clases seleccionado no permite garantizar la separabilidad entre estas, la información textural puede introducir información adicional que incremente la separabilidad de las clases.

En los últimos años se ha experimentado con el uso de información textural en teledetección, se trata del cálculo, píxel a píxel, de índices texturales que cuantifican los criterios texturales cualitativos tradicionales de la fotointerpretación.

La información textural suele extraerse del primer componente obtenido de un Análisis de Componentes Principales (ACP) realizado a las capas de reflectividad; para ello se ha utilizado la función semivariograma. Es la opción más utilizada en ciencias de la tierra y es muy intuitiva en cuanto a la naturaleza de sus parámetros.

En este trabajo se ha optado por incorporar dos índices texturales basados en el semivariograma:

$$\gamma = \frac{\sum_{i=1}^4 (b - b_i)^2}{8} \quad (7)$$

donde  $b$  hace referencia al valor en la celdilla analizada de la banda de que se trate,  $b_i$  hace referencia al valor en las cuatro celdillas vecinas (norte, sur, este y oeste).

Este índice se ha aplicado a dos bandas obtenidas a partir de los valores de reflectividad:

- Primer componente de un ACP con las bandas de reflectividad, que aparece relacionado casi siempre con un albedo medio ponderado de la imagen.
- Índice Normalizado de Vegetación (NDVI) que responde a la ecuación:

$$NDVI = \frac{\rho_4 - \rho_3}{\rho_4 + \rho_3} \quad (8)$$

Se ha implementado en un solo *script* todo el código necesario tanto para el cálculo de estas bandas de información textural como para el cálculo de las transformaciones logarítmica e inversa que se detalla en el apartado de análisis estadístico de las áreas de entrenamiento.

### Clasificación supervisada

Las especiales características de este proyecto imponen la necesidad de buscar áreas de entrenamiento que hayan permanecido con el mismo uso en los últimos 30 años, para ello se han utilizado mapas de uso de suelo e imágenes de tres fechas:

- Mapa de Cultivos y Aprovechamientos de 1977 y de 2001 .
- Ortofoto de 1986-87.
- Mapa del proyecto Corine-Land Cover de 1990 y 2000.
- Ortofoto de la Dirección General de Medio Natural de la CARM de 2008.

Se han buscado zonas con los mismos usos del suelo en todas las fechas, homogéneas en cuanto a topografía y que no presentasen discontinuidades.

El conjunto de clases seleccionado *a priori* aparece representado cartográficamente en la figura 2, un total de 550 áreas de entrenamiento que suman una superficie de 19.282 Hectáreas.



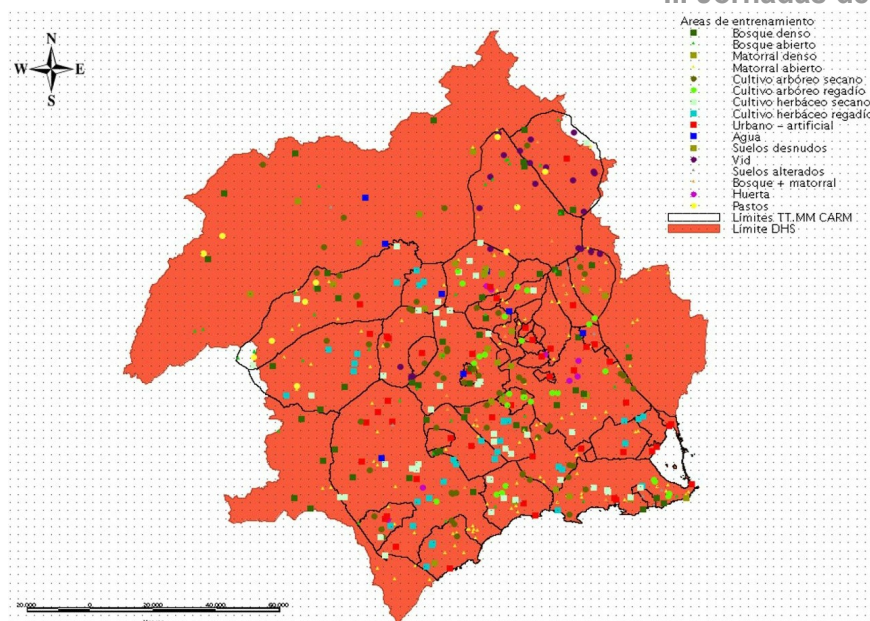


Figura 2: distribución de las áreas de entrenamiento en el área de trabajo.

Para esta labor ha sido especialmente útil el uso de el SIG Kosmo V 1.2 ya que es un programa muy bien planteado para la visualización de cartografía en formato vectorial y raster simultáneamente. Hay que destacar la capacidad y estabilidad del software para cargar servicios WMS, fundamentales en esta fase del trabajo.

## Análisis estadístico de las áreas de entrenamiento

### Test de normalidad y eliminación de outliers

La necesidad de cumplir las exigencias de un método de clasificación paramétrico como es la máxima verosimilitud, obliga a una comprobación exhaustiva de su cumplimiento, en particular la normalidad de las variables empleadas, en este caso las bandas corregidas atmosféricamente y la capa de semivariograma de albedo y NDVI.

En este sentido, y con carácter exploratorio, se han obtenido gráficos de normalidad para cada combinación de los valores de las bandas con las clases informacionales (figura 4). Estos gráficos se han obtenido también para transformaciones de variable logarítmica e inversa.

A partir de estos gráficos resulta fácil observar la presencia de *outliers* que pueden eliminarse recalculando el mapa de áreas de entrenamiento para eliminar de estas los píxeles con dichos valores.

Una vez eliminados los *outliers* se ha determinado, a través del análisis de los gráficos, la transformación más adecuada para cada banda. En cada banda deberá buscarse una transformación de compromiso entre todas las clases. Estos procedimientos se han implementado en lenguaje R.

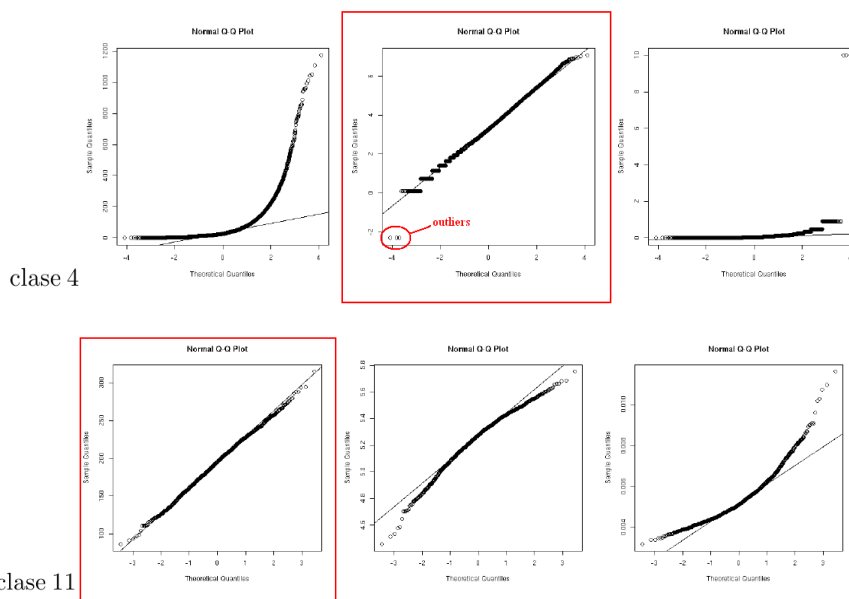


Figura 3: ejemplos de test de normalidad a las bandas originales (izquierda), transformación logarítmica (centro) e inversa (derecha). Además también se muestra la localización de outliers.

### Definición de clases

Una vez asegurada la normalidad de las bandas con que se va a trabajar y la eliminación de *outliers* en las áreas de entrenamiento; el siguiente paso ha sido la estandarización de las variables que resulta imprescindible para utilizar métodos de clasificación basados en *cluster*.

Estas variables estandarizadas se agrupan en un *group* de GRASS, se trata de estructuras que integran diversas bandas a las que va a realizarse un tratamiento conjunto.

Finalmente se crean los siguiente grupos de clases:

- Clasificación supervisada para caracterizar cada una de las clases definidas en la leyenda que acompaña a la figura 4 y cada uno de los polígonos que forman el conjunto de áreas de entrenamiento.
- Clasificación no supervisada de toda la imagen, el objetivo es descubrir si un método automático puede identificar en la imagen clases espectrales significativamente diferentes de las clases informacionales establecidas *a priori*.

### Integración de las clasificaciones supervisada y no supervisada

La integración de la clasificación supervisada y no supervisada se hace combinando categorías de ambos métodos empleando un algoritmo de agrupación jerárquico. Se utiliza como criterio de ordenación la divergencia entre las CE (clases espectrales) y las CI (clases informacionales). Si existe una CE con baja divergencia con todas la CI se considerará una clase significativa y se añadirá al conjunto de clases. Un método similar ha sido probado con buenos resultados en [5].

El índice de divergencia (ecuación 9) se ha utilizado para evaluar la similitud o disimilitud entre clases de manera que indica la posibilidad de distinguir una de otra [6] y [7]. Se trata de un test paramétrico que, por tanto, asume normalidad en las



variables que caracterizan las clases y será especialmente adecuado para comprobar la separabilidad entre clases empleando el método de máxima verosimilitud.

$$Div_{a,b} = \frac{Tr\{(V_a - V_b)(V_b^{-1} - V_a^{-1})\}}{2} + \frac{Tr\{(V_a^{-1} + V_b^{-1})(M_a - M_b)(M_a - M_b)^t\}}{2} \quad (9)$$

En la anterior ecuación  $M_a$  es el vector de medias de la clase  $a$ ,  $V_a$  es la correspondiente matriz de varianzas-covarianzas,  $Tr$  la función traza de la matriz,  $t$  indica la matriz transpuesta y  $^{-1}$  la matriz inversa.

### Asignación de píxeles a clases

Una vez obtenido un conjunto de clases óptimo y su caracterización estadística procede asignar cada uno de los píxeles presentes en la imagen a una de aquellas. Para esto se ha utilizado en primer lugar el algoritmo de máxima verosimilitud de manera que para cada clase se obtendrá la probabilidad de pertenencia a dicha clase.

La clasificación máximo verosímil calcula para cada clase  $A$  la probabilidad de que un píxel  $x$  pertenezca a ella a partir de la función de densidad multivariante [6]:

$$p(x/A) = (2\pi)^{-m/2} |V_A|^{-0.5} \exp\{-0.5(X - m_a)^t V_A^{-1} (X - m_A)\} \quad (11)$$

donde  $m$  indica el número de bandas,  $V_A$  es la matriz de varianzas-covarianzas para la clase  $A$ ;  $X$  el vector de valores del píxel  $x$  y  $m_A$  el vector de medias de la clase  $A$ .

### Validación. Error de clasificación

La tarea de validación de resultados se ha llevado a cabo analizando una serie de áreas test obtenidas del mismo modo que las áreas de entrenamiento pero que no fueron utilizadas para este. Esta forma de proceder permite obtener una estimación realista de los errores mientras la muestra de píxeles para la estimación del error sea lo suficientemente grande y representativa.

Para realizar este trabajo de forma semiautomática se han programado dos *scripts*, uno en BASH y otro en R en los cuales se implementan todos los cálculos necesarios para evaluar la calidad del resultado. Básicamente, estos son los que forman la matriz de confusión más los índices y coeficientes que se desprenden de ella [6]:

- Índice de Fiabilidad: es la relación entre el número de puntos correctamente asignados y el total.
- Coeficiente *kappa*: mide la diferencia entre el acuerdo mapa-realidad obtenido y el que cabría esperar simplemente por azar. Intenta delimitar el grado de ajuste debido sólo a la exactitud de la clasificación, prescindiendo del causado por factores aleatorios. Su fórmula de cálculo es:

$$\kappa = \frac{n \sum_{i=1}^n X_{ii} - \sum_{i=1}^n X_{i+} X_{+i}}{n^2 - \sum_{i=1}^n X_{i+} X_{+i}} \quad (12)$$

donde  $X_{ii}$  hace referencia al número de píxeles correctamente clasificados en la clase  $i$ ,  $X_{+i}$  es el total de píxeles pertenecientes a la clase  $i$  y  $X_{i+}$  es el total de píxeles clasificados en la clase  $i$ . Un valor de *kappa* igual a 1 indica un acuerdo pleno entre la realidad y el mapa, mientras un valor cercano a 0 sugiere que el acuerdo observado es puramente debido al azar.

- Errores de omisión ( $E_o$ ): aparecen en las celdillas marginales de las columnas de la matriz. Indican el número de píxeles que, perteneciendo a una determinada categoría, no fueron incluidos en ella. Para cada clase se calculan como:

$$E_{o,i} = \frac{X_{+i} - X_{ii}}{X_{+i}} \quad (13)$$

donde  $X_{+i}$  indica el marginal de la columna  $i$ , y  $X_{ii}$  la diagonal de dicha columna. Se refieren a una definición imperfecta de la categoría.

- Errores de comisión: de forma similar, las celdillas no diagonales de las filas expresan los errores de comisión; esto es, píxeles que se incluyeron en una determinada categoría perteneciendo realmente a otra:

$$E_{c,i} = \frac{X_{i+} - X_{ii}}{X_{i+}} \quad (14)$$

donde  $X_{i+}$  indica el marginal de la fila  $i$ , y  $X_{ii}$  la diagonal de dicha fila. Se refieren a una delimitación excesivamente amplia de la categoría.

- Exactitud del usuario: es un valor que está en relación inversa con los errores de comisión.

$$F_{u,i} = \frac{X_{ii}}{X_{i+}} \quad (15)$$

- Exactitud del productor: es un valor que está en relación inversa con los errores de omisión:

$$F_{p,i} = \frac{X_{ii}}{X_{+i}} \quad (16)$$

Con este tipo de análisis, se ha obtenido, no sólo una caracterización del error cometido, sino también una medida sobre la adecuación de la clasificación considerada a la realidad y de los parámetros utilizados para caracterizarlas. Puede por tanto utilizarse para definir un nuevo conjunto de clases para realizar una clasificación.

## RESULTADOS

El resumen de los resultados obtenidos en la clasificación se puede observar en la tabla 3. La mayoría de las imágenes están clasificadas con un Índice de Fiabilidad y un Coeficiente *Kappa* superior a 0,61.

A la vista de estos resultados se puede afirmar que son algo inferiores a los valores que se suelen obtener en este tipo de estudios. Concretamente, si los comparamos con un estudio realizado por Graeme G. Wilkinson en 2005 [8] en el que se analizaban más de 500 experimentos, los valores obtenidos en nuestro trabajo son un 18% inferiores para el Índice de fiabilidad y un 11,43% para el Coeficiente *Kappa*. En la figura 4 pueden observarse algunos de los resultados de la clasificación. Finalmente se ha calculado un total de 32 mapas de coberturas del suelo.

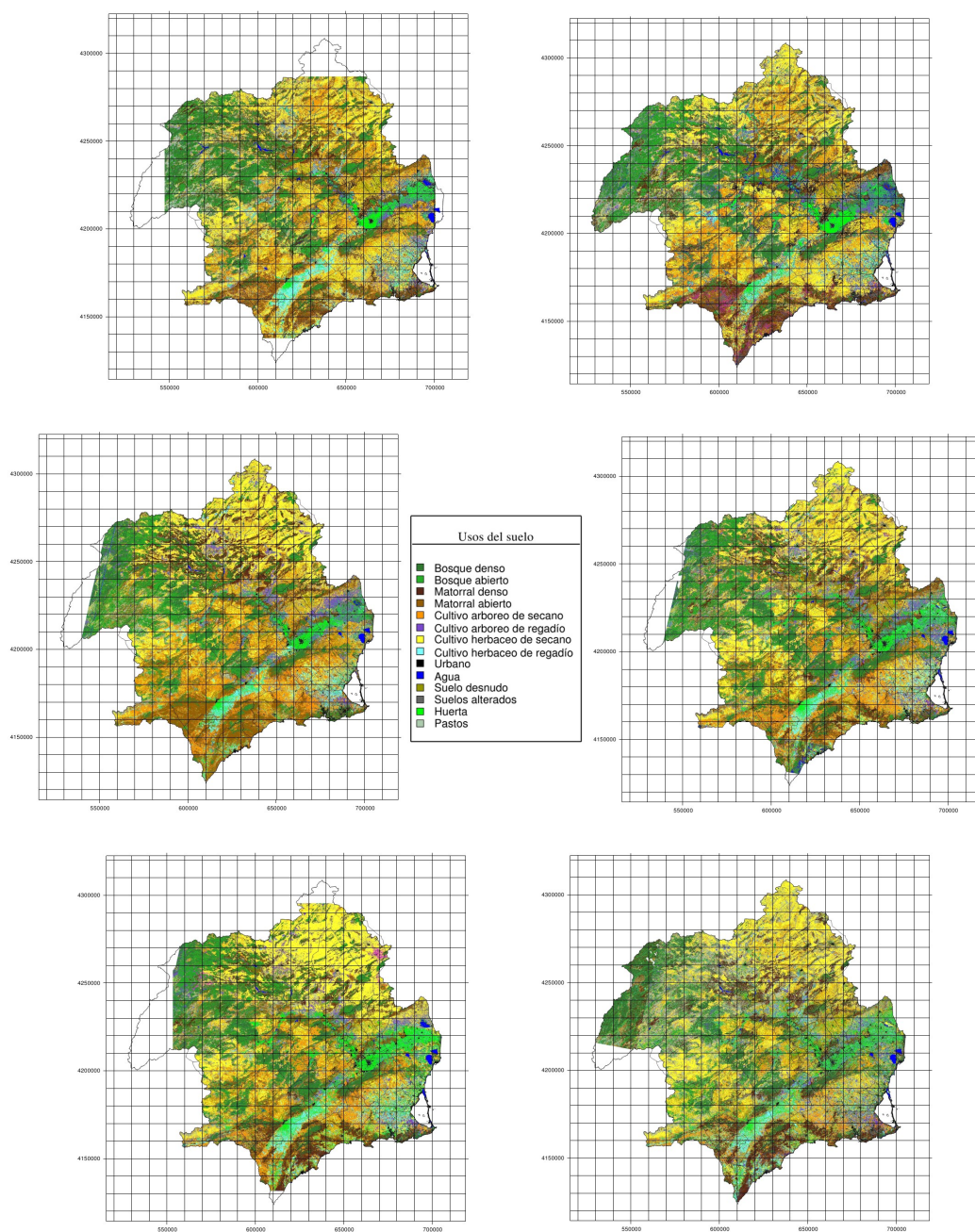


Figura 4: resultados de la clasificación, mapas de cobertura del suelo de los años 1976, 1982, 1988 y 1994, 2000 y 2006 (de izquierda a derecha y de arriba a abajo)

Tabla 3: Distribución de las medidas de fiabilidad de los mapas de cobertura de suelo.

Intervalo	Índice de Fiabilidad	Coefficiente Kappa
< 0.5	0	7
0.51-0.60	13	8
0.61-0.70	15	16
> 0.70	3	0

## CONCLUSIONES

A partir de enero de 2009 se ha producido la liberación de las imágenes captadas por los satélites Landsat. Esto pone a disposición de los usuarios de teledetección una fuente de información realmente abundante (se captan imágenes de la Península Ibérica aproximadamente cada 16 días desde 1972) del satélite más utilizado en ciencias de la tierra, a precio de adquisición cero.

La cuestión, siendo así las cosas, es si se dispone de los procedimientos apropiados para el pre-procesamiento de las imágenes, cuando éste se tiene que hacer sobre un número elevado de escenas casi simultáneamente. Entendemos por pre-procesado la corrección geométrica, atmosférica y de iluminación, puesta a punto de las áreas de entrenamiento, transformación de bandas, cálculo de información textural y en general todo el procesamiento de las imágenes previo a la ejecución de los algoritmos de clasificación.

En este sentido pensamos que el uso de sistemas operativos GNU/Linux, el SIG GRASS y el software de análisis de datos R, todos ellos en modo comando, constituyen en la actualidad la mejor solución a la cuestión que se plantea líneas arriba. Por lo siguiente:

- Precio de adquisición del software.
- Flexibilidad del entorno de programación Shell de Bash, teniendo en cuenta que en el mismo *script* se llama a distintos módulos de software específico y diferente, añadiendo la posibilidad de redirigir salidas de comandos como entradas de los siguientes.
- Estabilidad del sistema frente a otros SO propietarios .
- Trazabilidad del flujo de trabajo, un concepto procedente de la producción de bienes que tiene una aplicación directa a los procesos que corren sobre sistemas informáticos. En la actualidad se dispone de archivos de fácil acceso en los que están registrados para cada uno de los procesos realizados sobre las imágenes el usuario, fecha, variables, parámetros, comandos, etc. Es decir, toda la información necesaria para evaluar la adecuación o no de los resultados obtenidos así como la búsqueda de errores humanos.
- Posibilidad de recalcular todas las capas de información en un tiempo sustancialmente inferior en comparación a un sistema basado en ventanas.
- La metodología desarrollada es válida para un proyecto de las mismas características llevado a cabo por cualquier otra institución.
- Código de fuente abierto, en determinadas fases del proyecto se ha modificado el código fuente de los comandos de GRASS (*i.maxlik2*) para adaptarlo a las necesidades específicas del proyecto. Solución inviable en la mayoría del software propietario.

La primera fase del proyecto “Evolución espacio-temporal de los usos del suelo en el ámbito de la DHS” del que este trabajo forma parte, ha sido llevada a cabo satisfactoriamente ya que se ha logrado desarrollar formalmente y escribir el código para una metodología unificada del pre-procesamiento de imágenes procedentes del satélite Landsat. Además se ha conseguido una primera versión de los mapas de cobertura del suelo del ámbito estudiado.

Los procedimientos han sido implementados en forma de *script* de manera que para un futuro inmediato se trata de ir aislando sistemáticamente las posibles fuentes de error en la clasificación.

De esta forma se espera que se consiga el objetivo planteado a medio plazo: poner en producción, listo para formar parte de la infraestructura de datos espaciales del INUAMA e IEA, una colección multitemporal de mapas de cobertura del suelo para la DHS, cuyas medidas de acierto/fiabilidad estén dentro de los márgenes aceptados por la comunidad científica.

## AGRADECIMEINTOS

El desarrollo de este trabajo ha sido posible gracias al Proyecto de Investigación denominado “Evolución de los usos del suelo en la Cuenca del Segura mediante técnicas de Teledetección” financiado por la Fundación Instituto Euromediterráneo del Agua (IEA), dentro del convenio específico de colaboración entre el Instituto del Agua y Medio Ambiente (INUAMA) de la Universidad de Murcia y el IEA para la realización de actividades conjuntas de investigación en recursos hídricos.

## REFERENCIAS

- ♦ [1] RIAÑO, D; SALAS, J; CHUVIECO, E. (2000), “Corrección atmosférica y topográfica, información multitemporal y auxiliar aplicadas a la cartografía de modelos de combustibles con imágenes landsat TM”. *Tecnologías Geográficas para el Desarrollo Sostenible*, pp. 222-239. Universidad de Alcalá.
- ♦ [2] CHÁVEZ, P.S. (1988), “An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data”. *Remote Sensing of Environment* 24, pp. 459-479.
- ♦ [3] TEILLET, P.M.; GUINDON, B.; GOODENOUGH, D.G., (1982), “On the slope-aspect correction of multispectral scanner data”. *Canadian Journal of Remote Sensing* 58, pp. 84-106.
- ♦ [4] BERBEROGLU, S.; CURRAN, P.J.; LLOYD, C.D.; ATKINSON, P.M. (2007), “Texture classification of Mediterranean land cover”. *International Journal of Applied Earth Observation and Geoinformation* 9, pp. 322-334.
- ♦ [5] CHUVIECO, E.; CONGALTON, R.G. (1988), “Using cluster analysis to improve the selection of training statistics in classifying remotely sensed data”. *Photogrammetric engineering and remote sensing* 54(9), pp. 1275-1281.
- ♦ [6] CHUVIECO, E. (2002), Teledetección Ambiental, Barcelona, Ariel.
- ♦ [7] THOMAS, I.L.; CHING, N.P.; BENNING, V.M.; D’AGUANNO, J.A. (1987), “A review of multi-channel indices of class separability”. *International Journal of Remote Sensing* 8(3), pp. 331-350.
- ♦ [8] WILKINSON, G.G. (2005), “Results and implications of a study of fifteen years of satellite image classification experiments”. *IEEE Transactions on Geoscience and Remote Sensing* 43(3), pp. 433-439.